EXHIBIT F

Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer

Roger Heim and Roger Y. Tsien

Background: Variants of the green fluorescent protein (GFP) with different colors would be very useful for simultaneous comparisons of multiple protein fates, developmental lineages and gene expression levels. The simplest way to shift the emission color of GFP is to substitute histidine or tryptophan for the tyrosine in the chromophore, but such blue-shifted point mutants are only dimly fluorescent. The longest wavelengths previously reported for the excitation and emission peaks of GFP mutants are 488 and 511 nm, respectively.

Results: Additional substitutions, mainly in residues 145–163, have improved the brightness of the blue-shifted GFP mutants with histidine and tryptophan in place of tyrosine 66. Separate mutations have pushed the excitation and emission peaks of the most red-shifted mutant to 504 and 514 nm, respectively. At least three different colors of GFP mutants can now be cleanly distinguished from each other under the microscope, using appropriate filter sets. A fusion protein consisting of linked blue- and green-fluorescent proteins exhibits fluorescence resonance energy transfer, which is disrupted by proteolytic cleavage of the linker between the two domains.

Conclusions: Our results demonstrate that the production of more and better GFP variants is possible and worthwhile. The production of such variants facilitates multicolor imaging of differential gene expression, protein localization or cell fate. Fusions between mutants of different colors may be useful substrates for the continuous *in situ* assay of proteases. Demonstration of energy transfer between GFP variants is an important step towards a general method for monitoring the mutual association of fusion proteins.

Address: Howard Hughes Medical Institute 0647 and Department of Pharmacology, University of California, San Diego, La Jolla, California 92093-0647, USA.

Correspondence to: Roger Y. Tsien
Requests for samples must be on letterhead.

Received: **18 September 1995**
Revised: **15 November 1995**
Accepted: **29 November 1995**

Current Biology 1996, **6**:178–182

Background

The green fluorescent protein (GFP) of the jellyfish *Aequorea victoria* has attracted tremendous interest as the first, and so far only, example of a cloned protein whose expression generates strong visible fluorescence without any additional cofactors [1,2]. The biggest absorbance peak of this 238 amino-acid protein is at 395 nm, with a smaller peak at 475 nm, and the emission maximum is at 508 nm when excited at 395 nm. Mutants of GFP with altered excitation and emission spectra are important for at least three reasons: firstly, to provide distinguishable markers to monitor multiple cellular events simultaneously; secondly, to serve as donors and acceptors for fluorescence resonance energy transfer (FRET); and thirdly, to illuminate the structure–function relationship of an intrinsically fascinating protein (for a review, see [3]).

Previously, we described the first GFP mutants with emission wavelengths markedly shifted from wild-type [4]. In these mutants, the crucial tyrosine at position 66 that becomes the core of the fluorophore was mutated either to histidine or tryptophan, resulting in major shifts of both the excitation and emission peaks to shorter wavelengths. However, the intensity of fluorescence of these point mutants rather low. Here, we describe additional mutations that increase the brightness of these 'blue fluorescent proteins' (BFPs), as well as a mutant with the longest excitation and emission wavelengths yet recorded for a GFP variant. Fusion proteins containing two mutants of appropriately shifted GFPs show the expected FRET.

Results

Three cDNAs encoding mutant GFPs were used for further study. In these mutants, the tyrosine residue at position 66 was mutated to histidine (Y66H) or tryptophan (Y66W) [4], or the serine at position 65 was mutated to threonine (S65T) [5]. These cDNAs were separately further mutagenized using the polymerase chain reaction (PCR; see Materials and methods) and then transformed into *Escherichia coli*; the colonies were visually screened for unusual intensities or colors. Isolation, spectral characterization (Table 1 and Fig. 1) and DNA sequencing yielded three useful variants.

The double mutant Y66H/Y145F had almost the same wavelengths as the single mutant Y66H, but almost twice the brightness — due mainly to a higher quantum efficiency of fluorescence. This double mutant also

Table 1

Fluorescence properties of different GFP mutants.

Clone	Mutations	Excitation max (nm)	Emission max (nm)	Extinction coefficient ($M^{-1}cm^{-1}$)	Quantum yield
Wild-type	None	395 (475)	508	21 000 (7 150)	0.77
P4	Y66H	383	447	13 500	0.21
P4-3	Y66H Y145F	381	445	14 000	0.38
W7	Y66W N146I M153T V163A N212K	433 (453)	475 (501)	18 000 (17 100)	0.67
W2	Y66W I123V Y145H H148R M153T V163A N212K	432 (453)	480	10 000 (9 600)	0.72
S65T	S65T	489	511	39 200	0.66
P4-1	S65T M153A K238E	504 (396)	514	14 500 (8 600)	0.54

Numbers in brackets refer to the properties of minor peaks.

developed its fluorescence during overnight growth, whereas the single mutant required several days.

The first round of mutagenesis to increase the brightness of Y66W yielded a variant with three additional substitutions — M153T/V163A/N212K. The clone encoding this mutant was subjected to another round of mutagenesis, yielding two variants with further mutations, N146I and I123V/Y145H/H148R (Table 1). The quantum efficiency of these mutants was comparable to wild-type GFP. Although we have not dissected the relative contribution of the various substitutions, their clustering in positions

Figure 1

Fluorescence excitation and emission spectra of different GFP mutants. All spectra were normalized to a maximal value of 1. Each pair of excitation and emission spectra is depicted by a distinct line colour.

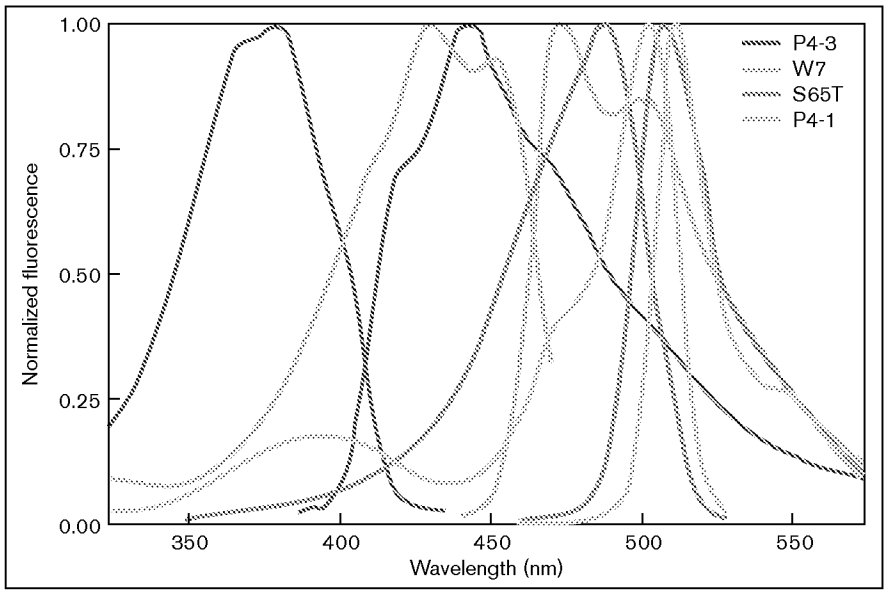
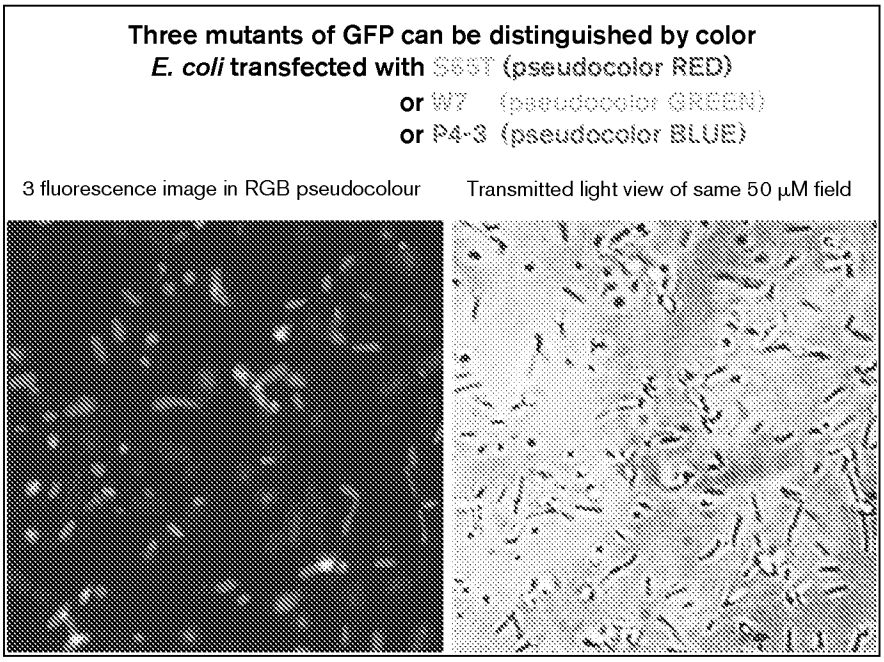


Figure 2



Three mutants of GFP are distinguishable by their excitation and emission spectra. *E. coli* were separately transfected with a cDNA encoding the P4-3, W7 or S65T mutant, mixed, and spotted onto a cover slip. Three fluorescence images were acquired using different filter sets that preferentially detected P4-3, W7 or S65T (see Materials and methods); to permit visual comparison, the three images thus obtained were displayed in blue, green and red pseudocolors, respectively, and overlaid with some small lateral displacements to adjust for imperfect registration of the raw images (left panel). A transmitted light view of the same field (50 μm square) is also shown (right panel). Most of the bacteria seen in the transmitted light image are sufficiently fluorescent to show up in the pseudocolor composite, and all of those show pure blue, green or red pseudocolor with no ambiguity, except where different cells touch each other. We had expected that matrix algebra might be necessary to correct for GFP signals spilling over into inappropriate detection channels, but this pseudocolor image was obtained without such mathematical manipulation. Please note that red is an arbitrary pseudocolor, not the physical color of S65T emission.

145–163 suggests that these residues lie relatively close to the chromophore, and that reductions in the size of their side chains might be compensating for the larger size of tryptophan compared with tyrosine.

Mutagenesis of the clone encoding S65T, to shift its wavelengths further to the red, yielded a variant with two further substitutions, M153A/K238E. This GFP variant had the longest-wavelength excitation maximum yet described, 504 nm, as compared with 490 nm for S65T (Table 1). Surprisingly, the emission peak hardly changed (514 nm as compared to 511 nm), so that the separation between the excitation and emission peaks (Stokes' shift) was extremely narrow, only 10 nm. This is one of the smallest values reported for any fluorophore in aqueous solution at room temperature. As in the Y66W series, M153 appears to be influential. We doubt that K238E is important, because this substitution of the carboxy-terminal amino acid was analyzed previously and found to be neutral [4].

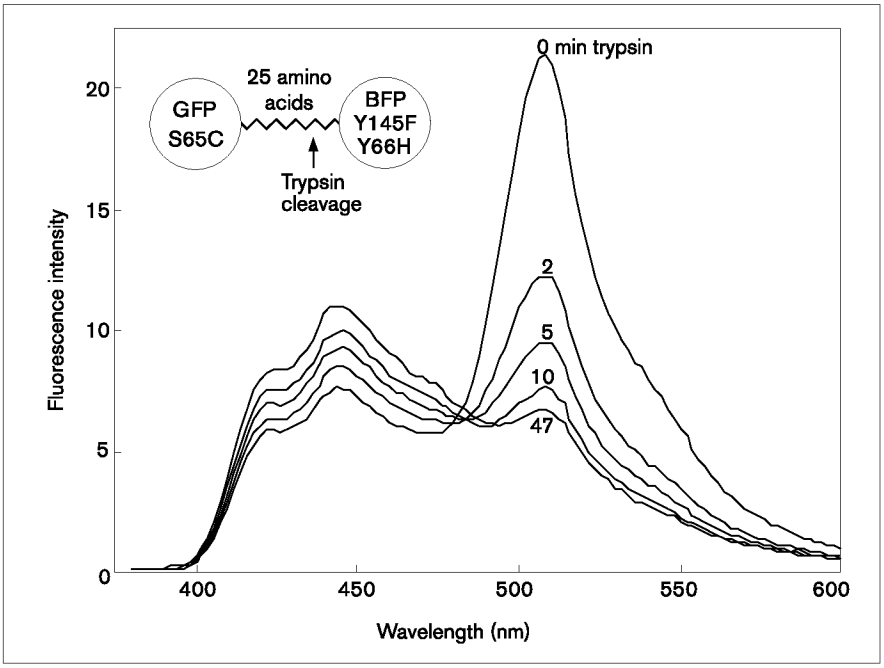
Among the many possible uses for GFPs with different colors, the most obvious may be the simultaneous monitoring of two or more promoter activities, protein localizations or cell fates. To test whether the currently available mutants could be distinguished by their excitation and emission wavelengths, *E. coli* expressing P4-3 (Y66H/Y145F), W7 (Y66W/N146I/M153T/V163A/N212K) or S65T were mixed, adhered to a cover slip, and sequentially imaged through a fluorescence microscope. Three filter cubes were chosen to optimize the brightness and distinguishability of each mutant: ultra-violet excitation/blue emission for P4-3, violet-blue excitation/cyan emission for W7, and cyan excitation/green emission for S65T. Each of the three images revealed a nonoverlapping subpopulation of the bacteria, such that a false-color superposition in blue, green and red showed pure primary colors, except where different cells were physically overlaid (Fig. 2). The three mutants are therefore quite distinguishable.

The availability of several differently colored mutants of GFP also opens the attractive possibility of fluorescence resonance energy transfer between them [6]. This quantum mechanical effect occurs when an excited fluorescent group is located within a few nanometres of a light-absorbing group. If the distance and mutual orientations are suitable, and the energy of the excited fluorophore coincides with the energy needed to excite the absorber, then the energy can be transferred. Such transfer is manifested by a reduction in the fluorescence intensity of the donor fluorophore, a reduction in the lifetime of its excited state and, if the acceptor is fluorescent, re-emission at its characteristic wavelengths.

In principle, the spatiotemporal dynamics of protein–protein interactions could be monitored in living cells if each partner were fluorescently labeled by fusion

Figure 3

Energy transfer between covalently linked BFP and GFP is abolished after cleavage with trypsin. The green GFP mutant S65C [5] was linked to the blue mutant Y66H/Y145F by expressing their fused cDNAs with a linker sequence that can be cleaved by trypsin or enterokinase. Excited at 368 nm, the uncleaved dimer emitted bright green light that gradually dimmed upon cleavage of the linker to separate the protein domains. As the cleavage by trypsin progressed (0, 2, 5, 10 and 47 min), more blue light was emitted. There was no further change after 47 min. Similar results were obtained with enterokinase cleavage (data not shown).



with a different color of GFP. Binding of the host proteins to each other would bring the appended GFP mutants within a few nanometres of each other, enabling FRET. Dissociation of the complex would disrupt FRET, because the efficiency, E, of FRET falls off rapidly with increasing distance, R, between the fluorophores, according to the equation $E = [1 + (R / R_0)^6]^{-1}$. The characteristic distance, R_0, at which FRET is 50 % efficient depends on the quantum yield of the donor, or shorter-wavelength, fluorophore, the extinction coefficient of the acceptor, the longer-wavelength fluorophore, and the overlap between the donor's emission spectrum and the acceptor's excitation spectrum. Calculated values of R_0 for Y66H/Y145F randomly oriented with respect to S65T and S65C are both 4.03 nm, because the slightly higher extinction coefficient of S65T compensates for its slightly longer emission wavelength [5].

To verify that FRET could indeed occur, the cDNAs encoding the donor and acceptor were concatenated and the proteins expressed. The cDNA construct also encoded a polyhistidine tag and a 25 residue cleavable spacer, the latter serving as a flexible linker between the two GFP domains. The emission spectrum of the intact fusion protein showed that FRET was fairly efficient (Fig. 3), because ultra-violet excitation caused substantial green emission from the acceptor S65C. After proteolytic cleavage of the spacer, which permitted the two domains to diffuse apart, the green emission almost completely disappeared, whereas the blue emission from the Y66H/Y145F was enhanced, because its excited state was no longer being quenched by the acceptor. Control experiments with the same proteolytic conditions applied to either GFP mutant alone showed no effect, arguing that the GFP domains *per se* are resistant to proteolysis, as is known to be the case for the native protein [7].

Discussion and conclusions

The further mutations to brighten the Y66H and Y66W variants of GFP will enhance the possibility of using two or three colors of fluorescent protein to track differential gene expression or protein fates. The chromophore in GFP is well buried inside the rest of the protein [3], so much of the dimness of the original point mutants was presumably due to steric mismatch between the substituted amino acid and the cavity optimized for tyrosine. The location of the beneficial mutations implies that residues 145–163 are close to the chromophore. The M153A/S65T mutant has the longest wavelengths and the smallest Stokes' shift of any known fluorescent protein that does not use a cofactor; unfortunately, it is not yet shifted far enough to make it easily separable from S65T in fluorescence microscopy. Nevertheless, the remaining mutants provide at least three distinguishable varieties of GFP as reporters for multiple gene expression events, protein localizations or cell fates.

The BFP–GFP fusion protein demonstrates that FRET can monitor the distance between GFP domains. Disruption of FRET between man-made chromophores

in a short synthetic peptide has been used before to assay proteases [8,9], but the use of GFP mutants as the fluorophores gives the unique possibility of replacing organic synthesis by molecular biology, and monitoring proteases *in situ* in living cells and organisms. FRET is also one of the few methods for imaging dynamic noncovalent protein–protein associations *in situ*. Yet larger values of R_0 would be desirable; the main requirements are further improvement in the quantum efficiency of the donor and the extinction coefficient of the acceptor.

Materials and methods

Mutagenesis

Random mutagenesis of the *gfp* cDNAs was done by increasing the error rate of the PCR with 0.1 mM $MnCl_2$ and unbalanced nucleotide concentrations. The templates used for PCR encoded the GFP mutants S65T, Y66H and Y66W. They had been cloned into the *Bam*H1 site of the expression vector pRSETB (Invitrogen), which includes a T7 promoter and a polyhistidine tag. The GFP coding region (shown in bold) was flanked by the following 5′ and 3′ sequences: 5′-G GAT CCC CCC GCT GAA TTC **ATG ... AAA TAA** TAA GGA TCC-3′. The 5′ primer for the mutagenic PCR was the T7 primer matching the vector sequence; the 3′ primer was 5′-GGT AAG CTT TTA TTT GTA TAG TTC ATC CAT GCC-3′, specific for the 3′ end of GFP, creating a *Hin*dIII restriction site next to the stop codon. Amplification was over 25 cycles (1 min at 94 °C, 1 min 52 °C, 1 min 72 °C) using AmpliTaq polymerase (Perkin Elmer). Four separate reactions were run in which the concentration of a different nucleotide was lowered from 200 μM to 50 μM. The PCR products were combined, digested with *Bam*HI and *Hin*dIII, and ligated to the pRSETB cut with *Bam*HI and *Hin*dIII. The ligation mixture was dialyzed against water, dried and subsequently transformed into the bacterial strain BL21(DE3) by electroporation (50 μl electrocompetent cells in 0.1 cm cuvettes, 1900 V, 200 ohm, 25 μF). Colonies on agar were visually screened for brightness as previously described [4]. Approximately 7000 colonies were examined in each successful round of mutagenesis, which we do not claim to be exhaustive. The selected clones were sequenced with the Sequenase version 2.0 kit from United States Biochemical.

Construction of GFP–BFP concatenation

The DNA of the GFP mutant S65C [5] was amplified by PCR (1 cycle 3 min 94 °C, 2 min 33 °C, 2 min 72 °C; 20 cycles 1 min 94 °C, 1 min 44 °C, 1 min 72 °C) with Pfu polymerase (Stratagene), using the primers 5′-AGA AAG GCT AGC AAA GGA GAA GAA C-3′ and 5′-T CAG TCT AGA TTT GTA TAG TTC ATC-3′ to create a *Nhe*I site and a (*Nhe*I compatible) *Xba*I site, and to eliminate the GFP stop codon. The restricted product was cloned in-frame into the *Nhe*I site of the construct pRSETB-Y66H/Y145F, between the polyhistidine tag and the enterokinase cleavage site. When translated this fusion gives the following sequence: MRGSHHHHHHGMA-(S2...GFP:S65C...K238)-SSMTGGQQMGRDLYDDDDKDPPAEF-(GFP:Y66H/Y145F).

Isolation of GFP

Cultures with freshly transformed cells were grown at 37 °C to an optical density of 0.8 at 600 nm, then induced with 0.4 mM isopropyl-thiogalactoside overnight at room temperature. Expression levels were roughly equivalent between mutants and are typical for the T7 expression system used. Cells were washed in PBS pH 7.4, resuspended in 50 mM Tris pH 8.0, 300 mM NaCl and lysed in a French press. The polyhistidine-tagged GFP proteins were purified from cleared lysates on nickel-chelate columns (Qiagen) using 100 mM imidazole in the above buffer to elute the protein. Samples used for proteolytic experiments were further purified by MonoQ FPLC to remove monomeric GFP. Protein concentrations were estimated with bicinchoninic acid (BCA kit from Pierce) using bovine serum albumin as a standard.

Proteolysis

Proteolytic cleavage of 10 μg of the GFP–BFP fusion protein was performed in 500 μl PBS pH 7.4 with 0.1 μg trypsin (Sigma, grade III), and emission spectra were recorded at different time intervals. Analogous cleavage experiments were done also with enterokinase (Sigma).

Spectroscopy

Excitation spectra were obtained by collecting emission at the respective peak wavelengths and were corrected by a Rhodamine B quantum counter. Emission spectra were likewise measured at the respective excitation peaks and were corrected using factors from the fluorometer manufacturer (Spex Industries, Edison, NJ). In cleavage experiments, emission spectra were recorded at excitation 368 nm. For measuring molar extinction coefficients, 20–30 μg of protein were used in 1 ml PBS pH 7.4. The values in Table I necessarily assume that the protein is homogeneous and properly folded; if this assumption is incorrect, the real extinction coefficients could be yet higher. Quantum yields of wild-type GFP, S65T and P4-1 mutants were estimated by comparison with fluorescein in 0.1 N NaOH as a standard of quantum yield 0.91 [10]. Mutants P4 and P4-3 were likewise compared to 9-aminoacridine in water (quantum yield 0.98). W2 and W7 were compared to both standards, which fortunately gave concordant results. For the experiment shown in Figure 2, *E. coli* were separately transfected with a cDNA encoding the P4-3, W7 or S65T mutant, mixed, and spotted on a cover slip, which had been pretreated by a 5 min exposure to 33 μM polylysine followed by a saline rinse. After 2 h for the bacteria to settle and adhere, three fluorescence images were acquired with a Photometrics (Tucson, AZ) cooled charge-coupled-device camera on a Zeiss Axiovert microscope with a 100×/1.3 numerical aperture Plan-Neofluar objective and 150 W xenon lamp. The three filter sets (all from Omega Optical, Brattleboro, VT) and exposure times were: 330WB80 excitation, 400DCLP02 dichroic, 440DF40 emission, 10 s; 433DF47 excitation, 455DRLP dichroic, 480DF30 emission, 1.2 s; and 488DF10, DR510LP, 535DF45, 0.8 s. In the filter names, three-digit and two-digit numbers indicate midpoint wavelength and bandwidth in nm. These three sets preferentially detect P4-3, W7 and S65T, respectively.

References

1. Prasher DC, Eckenrode VK, Ward WW, Prendergast FG, Cormier MJ: **Primary structure of the *Aequorea victoria* green fluorescent protein.** *Gene* 1992, **111**:229–233.
2. Chalfie M, Tu Y, Euskirchen G, Ward WW, Prasher DC: **Green fluorescent protein as a marker for gene expression.** *Science* 1994, **263**:802–805.
3. Cubitt AB, Heim R, Adams SR, Boyd AE, Gross LA, Tsien RY: **Understanding, using and improving green fluorescent protein.** *Trends Biochem Sci* 1995, **20**:448–455.
4. Heim R, Prasher DC, Tsien RY: **Wavelength mutations and post-translational autooxidation of green fluorescent protein.** *Proc Natl Acad Sci USA* 1994, **91**:12501–12504.
5. Heim R, Cubitt AB, Tsien RY: **Improved green fluorescence.** *Nature* 1995, **373**:663–664.
6. Tsien RY, Bacskai BJ, Adams SR: **FRET for studying intracellular signalling.** *Trends Cell Biol* 1993, **3**:242–245.
7. Ward WW, Prentice HJ, Roth AF, Cody CW, Reeves SC: **Spectral perturbations of the *Aequorea* green fluorescent protein.** *Photochem Photobiol* 1982, **35**:803–808.
8. Krafft GA, Wang GT: **Synthetic approaches to continuous assays of retroviral proteases.** *Methods Enzymol* 1994, **241**:70–86.
9. Knight CG: **Fluorimetric assays of proteolytic enzymes.** *Methods Enzymol* 1995, **248**:18–34.
10. Miller JN: *Standards in Fluorescence Spectrometry*. New York: Chapman and Hall; 1981.